DOCUMENT RESUME

ED 254 539 TM 850 117

AUTHOR Goldstein, Harvey; Nuttall, Desmond
TITLE Recent Developments in Assessment Procedures in England and Wales.
PUB DATE Apr 85
NOTE 9p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, IL, April 1-3, 1985).
PUB TYPE Reports - Descriptive (141) -- Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; Criterion Referenced Tests; *Educational Assessment; *Educational Change; Educational Testing; Evaluation Needs; *Foreign Countries; High Schools; Student Certification; Student Evaluation; *Testing Problems; *Testing Programs
IDENTIFIERS Certificate of Pre Vocational Education; *England; General Certificate of Education; *Wales

ABSTRACT

Focusing on technical issues, this paper critiques proposed changes in assessment procedures at the further educational level (ages 16 through 18) in England and Wales. Major structural changes are taking place at this educational level, partly because of large scale youth unemployment. The two current examination systems for the final year of cumpulsory education, the Ordinary (O) level of the General Certificate of Education (GCE) and the Certificate of Education (CSE), are now being combined in a General Certificate of Secondary Education (GCSE) with courses starting in 1986. A new 17-plus exam, the Certificate of Pre-Vocational Education (CPVE) is also being introduced. In addition two major initiatives are taking place in the schools: (1) student profiles, or records of achievement; and (2) graded or graduated assessments. All these innovations are characterized by a commitment to development and implementation without adequate evaluation. This lack of concern for curricular and organizational consequences or major technical problems (such as criterion referencing, score equating, and reliability) reflects a widespread disillusion with the academically divisive and unreliable current examination system (BS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED254539

Recent Developments in assessment procedures in England and Wales

The 5-18 year pre university system in England and Wales is usefully divided into Primary (5-11), compulsory Secondary(11-16) and Further(16-18). Apart from private schools(with about 6% of children), the only institutions which do not neatly fit this pattern are middle (8-13) schools of which there are rather few, and special schools. At the further education level, about half the full time students are in school sixth forms or colleges and half in so called further education institutions. In the public sector, schools and further education institutions are maintained by the local education authorities. The latter receive something over half their income from Central Govt. and the rest from local taxes, the level of which is coming under increased Central Govt. control.

At present, the major changes in the structure of the system are taking place at the further education level , partly because of large scale youth unemployment. Traditional school sixth forms are tending to be grouped together across schools and a less sharp distinction is apparent between these and further education institutions. The latter and to some extent schools themselves, are collaborating with the Manpower Services Commission(MSC) vocational youth trainig schemes and introducing pre vocational courses on a large scale to cater for the growing number of unemployed 16 and 17 yerr olds. This is also having an effect on courses for 14 and 15 year olds. In the next section we will outline the current assessment scene, then describe proposals for change and finally examine some of the technical and resource issues.

Current Systems of Examining

At 16+, in the final year of compulsory schooling, there are two systems of examinations. One is the Ordinary(O) level of the General Certificate of Education(GCE). It is the older and consists of separate subject exams and is conducted by 8 boards. Some of these, like the London and Oxford Board have traditional strong University links, while others, such as the Associated Examining Board are independent. All, in principle, have national coverage and are essentially commercial organisations competing

BEST COPY

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Goldstein

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

with each other for a share of the market, albeit in close cooperation with LEA's, schools, teacher unions and other educational interests. These boards also administer the Advanced level of the GCE exam, a single subject exam in which most candidates take three subjects, and which is principally used as a university entrance qualification and selection device. Under economic pressure, the boards have also recently been diversifying their activities, for example providing basic testing materials for schools and industry. In addition, some boards have become heavily involved in the graded assessment and profiling movements which we describe below. The other system at 16+ is the Certificate of Secondary Education(CSE), introduced in 1965, administered by regional boards and designed for the approximate 40-80 percentile ability range. A grade 1 in the CSE is deemed to be 'equivalent' to a C grade(out of grades A-E) O level and a grade 4 is meant to represent 'average' attainment in a subject. The most important innovation in the CSE is the use of so called 'mode 3' which is where the exam is set and marked by a school or schools with the exam board adopting a moderating role. A government body, the Secondary Examinations Council, has responsibility for monitoring and promoting school examination systems. In the further education system, the principal examining systems are those provided by the City and Guilds of London Institute and the Business and Technician Education Council(BTEC). The former is responsible for a series of nationally accredited diplomas and certificates at craft level, and the latter for the maintenance of a series of higher level syllabuses which are locally examined with centralised moderation . In addition, there are other bodies, such as the Royal Society of Arts, which offer more specialised, and often innovatory, assessments. In the further education sector, a government body, the Further Education Unit, has an overall brief including that of monitoring and promoting assessment, but without control.

Testing in Schools

A crucial feature of the examining systems described above is their link with a prescribed syllabus. Standardised tests, on the other hand, do not typically share this property and their use in the English and Welsh system, with one notable exception, has been confined to non certification . The exception is the 11+ selection procedure, still existing in parts of some 15% of LEA's, which has relied heavily on standardised tests for

BEST COPY AVAILABLE

the purposes of selecting children for Grammar or secondary Modern schools. Apart from this, and coinciding with the gradual removal of the 11+ , there has been an increasing use of tests by LEA's at the primary school level. While some of this has coincided with an increased interest in screening for special educational needs, the majority has occurred for the purpose of monitoring and comparing the attainments of groups of children, especially over time. In practice, however, little use is made of the results of this testing. It is not at present used for teacher assessment or generally to assist in resource allocation procedures(Gipps et al., 1983). Nevertheless, there is a widespread use of standardised tests - about 80% of LEA's in 1981 had some kind of testing programme, and there seems to be an increasing interest in using the results for such purposes.

Proposals for Change

a) Examinations:

In schools, the major change is the amalgamation of the GCE and CSE exams into a combined General Certificate of Secondary Education(GCSE). Courses start in 1986 and current activity includes the preparation of criteria, syllabuses and a very basic training of teachers in the new procedures. The major pressure for this rationalisation has been the difficulty and divisiveness of maintaining two separate systems which are meant to have a specific relationship with each other. Under the guidance of central govt. it has been decided that the new exam will be administered by 5 new examining groups formed by bringing together existing GCE and CSE boards. It has also been decided that the design of the exams will follow Criterion Reference procedures in the form of 'grade related criteria', to which we return below. The govt. has taken much greater control over exam syllabus content and assessment techniques, for example by insisting that practical science and oral english be assessed.

The other major innovation is the introduction of a new 17+ exam, the Certificate of Pre Vocational Education(CPVE). This is to be administered by a joint board of the City and Guilds and BTEC and is intended to cater for those who stay in education beyond the compulsory leaving age who wish to have qualifications relevant to a wide range of possible employments. Both these innovations clearly have implications for curriculum and it is not yet clear how they will relate to each other, especially since a new 'Technical and Vocational Education Initiative(TVEI) is being introduced

BEST COPY AVAILABLE

into the school system for 14-18 year olds and for whom the CPVE would seem possibly a more appropriate exam than the GCSE or A levels. The CPVE, moreover, is principally envisaged as an exam to be taken in further education institutions rather than schools, and its introduction into schools is a further complication. The CPVE is intended to consist of partly a centralised assessment and partly a locally completed profile, and we shall return to this below. In addition to these developments, there will continue to be specialised vocational and non vocational exams administered by bodies traditionally active in this area, and including those bodies seeking new business such as the old GCE exam boards who will still retain their separate existence following the introduction of the GCSE.

b) Non exam assesment:

At the same time as the introduction of these new exams, two major initiatives are taking place in schools, and one of them in Further education also.The first is the widespread interest in 'profiles' or 'records of achievement' which have been the subject of experimentation at the further education level for some years and are now being developed in several LEA's with govt. support. In further education they are likely to be a major component of the CPVE and will consist of an extensive report of a student's achievements in academic, practical and possibly social and personal areas. It is envisaged that they will be completed collaboratively by students and lecturers or teachers, often in the form of a series of standard descriptions exemplifying stages of skill acquisition or task mastery. The development at the school level is less far advanced but is likely to follow similar lines, though more dependent upon descriptions in continuous prose and less on ticks on checklists. There are presently no plans to incorporate profiles into the GCSE. There are problems of comparability between single subject exams like GCSE and records of achievement, although at least one exam board is planning to issue a 16+ certificate which will contain profile information along with public exam passes. The other major development is the development of so called graded or graduated assessments. Pioneered by modern language teachers, these are now under considerable development with govt., exam board and charitable foundation support. They were advocated strongly in the Cockcroft report in 1983 on mathematics teaching and the chair of the cttee. producing that report is currently the chair and chief executive of the Secondary Exams

BEST COPY AVAILABLE

Council. The basic idea of graded assessment is that students are assessed either by a formal test or, for example, on a teacher checklist, only when they are 'ready'. Up to 11 levels in the secondary school have been suggested, and in most cases specific equivalences are to be incorporated with GCSE grades.One of the principal motivations has been the claim that frequent rewards, each achieved with a high probability of success, will promote learning, especially among 'low attainers' on the principles of mastery learning. Analogies with music and swimming exams are often adduced. Only one of the current major projects in this area is a feasibility study, and that is specifically aimed at those who would be unlikely to take GCSE and hence is not concerned with linking to GCSE. The remaining projects appear to be commited to the introduction of graded assessment and there seems little doubt that this will come to have important implications for the curriculum.

Technical Aspects

The commitment to introducing the GCSE involves two novel aspects. The first is the differentiated papers for a single subject which require that different papers are equated to a common mark or grade scale. This is a case of 'vertical equating' where it has been shown (Morris, 1982) that this is strictly not possible where the different tests or exams cover different ability ranges. The only possible solution is to make the assumption of strict unidimensionality, which clearly is innapropriate in the present case. It is not a feature of the present concern with assessment that funding agencies have devoted much effort or concern with such basic technical problems. The other issue is that of criterion referencing. Criterion referenced tests were introduced and have been developed in the context of providing statements about mastery in well defined domains. The development of criterion referenced tests typically involves extensive piloting, and there seems to be little available concerning the substitution or equating of one such test to another.Indeed, the notion of test equating, which is essentially a 'norm referencing' process would seem to be somewhat inappropriate in the case of criterion referencing, and at the very best would be a necessary but not sufficient condition for substitution to be possible. Yet in the case of public exams the equating requirement is paramount since the exams have to be changed regularly and there is a strong demand for the maintenance of 'standards'. The other problem is that to specify

BEST COPY AVAILABLE

'criteria' for each grade of the exam means describing the mastery of very wide and complex domains which is almost certainly impossible to do, without disaggregating the exam into an impossibly large number of components. Nevertheless, the SEC, under directives from central govt., is attempting to achieve precisely this.

The graded test developments share some of the same problems of the GCSE, not the least because of the explicit link between them. While it may be possible to provide a reasonable criterion referencing at the lower levels of the tests, this will become more difficult at higher levels. There is also the problem of equating - not merely because of their attempt to be criterion referenced, but also because they can theoretically be taken 'when a student is ready' and this demands a certain amount of continual updating to preserve comparability. Comparability also would be required across schools and teachers for any localised system, which is felt desirable by many. Since the probability of passing an assessment has to be high, in order not to discourage students, this would seem to imply at least some kind of pre-assessment by teachers of a student's chances of success and effectively leading to a kind of shadow assessment system, which may well come to obviate the formal assessment itself, and at worst lead to an enormous effort centered around the graded assessment system. Again, the curriculum implications of this have hardly been discussed. There are also implications for school organisation, since the progression of students through a school will tend to depend on their assessments rather than their age. Perhaps most important of all, however is the linking of grades to GCSE grades, except in the 'low attainers' case mentioned above. One of the original aims of the graded assessment movement was the short term motivation towards learning freed of the long term dominance of school leaving exams. Yet by tying the system directly to the school leaving exam it would seem likely that the opposite could occur, and that the constraints of the school leaving exam will be manifest more clearly throughout the school. Ironically, the development of graded assessment is seen by some of the exam boards involved as a means of diversification whereby they can use their accumulated expertise in school leaving exams to help devise these assessments. Graded assessments are perhaps the most important current example of the use of assessment to determine, implicitly, a new curriculum. Thus, the

BEST COPY AVAILABLE

concept of a hierarchy of levels carries clear implications for curriculum structure, as of course does the content of the tests themselves. There has been very little discussion and no apparent major concern with such curriculum implications on the part of the funding bodies.
Profiling in many ways is a more radical departure as far as schools and further education institutions are concerned than either graded tests or the exam innovations. The idea of recording affective, attitudinal as well as cognitive achievement raises a very wide range of issues, both social and educational. From the technical point of view the principal issues have to do with criterion referencing, comparability and reliability.The basic item of a profile is a verbal description of achievement. The intention is that this should describe an 'underlying' skill or attribute and not merely record instances of specific achievement. The difficulty is to achieve any kind of comparability when the nature of so called underlying competencies is notoriously difficult to define. Yet the more precise the descriptions becomes, the more they begin to resemble standardised tests, which would be self defeating. The assumption that it is possible accurately to describe competencies in this way seems to have hardly been questioned. In those profile systems which incorporate a centralised exam and local descriptions of achievement, there seems to be a danger that the central component will become dominant and used for selection etc. with the local assessment relegated to a kind of extended character reference. Many of those concerned with the social aspects have referred to the possibilities of social control inherent in this.The other major problem with profiles is that of reliability. Limited experience so far suggests that judgemental reliability is low and that interpretations of grading statements can vary widely. Furthermore, existing systems expect a single description to be chosen, whereas in reality a probabilistic assignment of a students performance across categories is more realistic, and failure to accommodate this will tend to increase unreliability. Finally there are the problems of scaling and weighting components of a profile. Even if a grading system for each attribute is only ordinal, any user of the system will inevitably have to combine the separate ratings in forming overall judgements. If no explicit advice is given this might be taken as an invitation to form an equally weighted composite, and in general it will not be realistic for the producer of a profile to avoid giving guidance on

BEST COPY AVAILABLE

this issue. It seems that there has been little attention to the aggregation issue in this form. We note some of the common strands in the current series of innovations. They all are characterised by a commitment to development and implementation without adequate evaluation. They rely heavily on the rhetoric of criterion referencing as opposed to norm referencing, as a legitimating device. They show little concern either with curriculum innovation or technical problems. They are funded largely from central govt. and presumably are seen by central govt. as a means of influencing the education system in line with a clear centralising policy which originated in the 1970's. They also occur at a time of cutbacks in educational spending and general economic difficulty, yet can only have the function of imposing a further burden on a system under pressure. This fact makes it all the more important to argue for careful evaluation and thought before implementation, although it now seems fairly clear that only after large scale implementation will the problems become widely visible. Finally, it seems to us that this lack of desire for critical evaluation reflects a widespread disillusion with the academically divisive and unreliable current examining system.

Harvey Goldstein, Institute of Education,
University of London

Desmond Nuttall, Open University

References

Morris, C. N. On the foundations of test equating. In: Test Equating, edited by Holland and Rubin, Academic Press, New York.

Gipps, C., Steadman, S., Blackstone, T., and Steirer, B. (1983). Testing Children. Heinemann, London.